Sequencing β-Casein C: Isolation of a Large Fragment After Cleavage of Thioltrifluoroacetylated β-Casein C

M. L. GROVES, W. G. GORDON, R. GREENBERG, R. F. PETERSON
Eastern Regional Research Center[1]
Philadelphia, PA 19118
R. JENNESS
Department of Biochemistry
College of Biological Sciences
University of Minnesota
St. Paul 55101

Abstract

A large fragment of β-casein C consisting of residues 26 to 184 was isolated by trypsin cleavage of lysine derivatized protein. The first 26 amino acid residues of the fragment are identical to those of the corresponding region in β-casein A^2 with the exception of a lysine/glutamic acid substitution and the absence of phosphorus on a serine residue phosphorylated in β-A^2. The apparent absence of γ-casein in milks containing β-casein C is also discussed.

Introduction

In earlier papers we presented evidence which suggested that γ-, TS-, R-, and S-caseins are fragments of β-casein. Thus, γ-casein corresponds to β-casein after 28 N-terminal amino acid residues are removed while TS-, R-, and S-caseins, depending on the polymorph, represent the C-terminal half of the corresponding β-casein type. One hypothesis for the origin of these components is that γ-, TS-, R-, and S-caseins result from a specific proteolytic cleavage of β-casein at specific sites (2, 5, 6, 7).

In milks from western breeds of cattle five polymorphs of β-casein designated A^1, A^2, A^3, B, and C have been identified. In milks containing each variant a corresponding γ-casein occurs except that γ-casein has never been found in milk typed β-casein C (8). Comparative amino acid analysis indicates that β-C differs from the other polymorphs by a lys/glu substitution (6). It also has only four phosphorus atoms compared to five for each of the other β-caseins. Because of these differences, it appears that sequence studies on β-C and comparison with sequences of the other β-casein polymorphs might provide an explanation of why no γ-casein C is found.

Ribadeau Dumas et al. (11) determined the primary sequence of β-casein A^2 and placed the phosphoserines at residues 15, 17, 18, 19, and 35. By examining appropriate peptides from the other β-casein polymorphs, Grosclaude et al. (4) provided evidence that residues 15, 17, 18, and 19 are phosphoserines in all variants, including the comparatively rare β-C. If this is so, then residue 35 should be serine (not phosphoserine) since β-C contains only four phosphorus atoms.

The availability of a sample of β-casein C isolated from the milk of a cow homozygous for β-C made it possible for us to sequence 22 residues from the N-terminal end of the molecule and to locate directly the four phosphoserine residues in the specified positions. To examine the amino acids in the neighborhood of residue 35, we isolated a large fragment of β-casein C starting with residue 26 and determined its N-terminal sequence.

While this work was in progress Grosclaude et al. (4) established the amino acid sequence around residue 35 for β-casein C. This communication confirms the amino acid sequence from residues 33 to 38 for β-C as reported by Grosclaude et al. (4) and shows that the reaction of ethyl thioltrifluoroacetate (ET-TFA) with β-C effectively modifies lysine residues so that enzymatic cleavage of the protein results in a large fragment of β-casein which can be used for sequencing.

Experimental Procedure

β-Casein C. β-Casein C was isolated from the milk of a Guernsey cow homozygous for β-C as described previously (5).

Isolation of a large fragment of β-casein C. β-Casein C contains 4 arginine and 12 lysine residues (6). To mask the lysine residues so that only the arginine bonds were sensitive to trypsin, ET-TFA was used according to the method of Goldberger (1). A suspension of

Received July 25, 1974.

[1] Agricultural Research Service, U.S. Department of Agriculture.

β-casein C, .88 g in 70 ml water, was placed in an autotitrator and dissolved by titrating to pH 9.0 with 5N NaOH. Then 2.5 ml ET-TFA were added to the solution, and the pH was maintained at 9.0. In 10 min another 2.5 ml aliquot of ET-TFA was added, and the solution was stirred for 1 h with pH at 9.0. Hydrochloric acid was added to the protein solution to give a pH of 4.6. After centrifugation, the precipitate was washed with water, alcohol, acetone, and air dried. Yield was .86 g of TFA β-C.

The TFA β-C was suspended in 21 ml water and dissolved by the addition of 2 N NaOH to a pH of 8.0, 25 C. A solution of trypsin (chymotrypsin inhibited, 8.2 mg in 2 ml H_2O) was added and the pH maintained at 8.0. After 1 h alkali consumption had leveled off, and the reaction was stopped by heating the solution 5 min at 100 C. The digest was then subjected to gel filtration at 3 C with Biogel P-60[2] (50 to 150 mesh) in a column, 2 by 93 cm, equilibrated with .05 M ammonium bicarbonate pH 8.0. The major peak containing the large fragment (TFA β-C_{Fg})[3] eluted near the void volume, was followed by a second minor peak which consisted of smaller fragments. The recovered TFA β-C_{Fg} amounted to .60 g. For removal of TFA from TFA β-C_{Fg}, 124 mg were dissolved in 2.0 ml 1 M piperidine and held at 3 C for 2 h. It was then poured into 3 ml of 1 M acetic acid previously cooled to 3 C. β-C_{Fg} was separated from the reagents by gel filtration on a Biogel P-4 column, 1 × 30 cm, equilibrated with .5 M acetic acid. β-C_{Fg} was recovered by lyophilization and amounted to 110 mg.

β-C_{Fg} is soluble at pH 6 but on titrating with NaOH the solution becomes increasingly turbid, especially between pH 7.0 to 7.7; at pH 8.3 turbidity decreases considerably. Cooling the mixture from room temperature to 3 C also increases the solubility of β-C_{Fg}. β-C_{Fg} (110 mg + 4 ml .005 M sodium phosphate pH 8.3) was dialyzed at 3 C against 2 liters phosphate buffer 4 h, then dialyzed against 2 liters fresh buffer solution overnight. This gave a clear protein solution which was applied to a microgranular DEAE cellulose column, 2 × 73 cm, equilibrated with .005 M sodium phosphate, pH 8.3.

Most of the β-C_{Fg} was eluted with the starting buffer in three fractions designated β-C_{Fg} I, β-C_{Fg} II, and β-C_{Fg} III in order of their elution sequence. β-C_{Fg} II was visualized as a shoulder on the falling side of the large first peak and β-C_{Fg} III was eluted in a smaller fraction somewhat later. β-C_{Fg} I and II were resolved by rechromatography on DEAE cellulose. β-C_{Fg} I was present in largest amounts although quantitation of these fractions was not attempted because only the center portion of each fraction was used for rechromatography.

Amino acid composition. Procedures of Moore and Stein (10) were used for automated amino acid analysis on samples hydrolyzed 24 h.

Sequence analysis. Sequence analysis was carried out on a Beckman 890 C sequencer, using a Quadrol double cleavage program (042772). The phenylthiohydantoin (PTH) amino acids were identified by gas-liquid and/or thin-layer chromatography and confirmed in many instances by HI (12) hydrolysis to the amino acid. The aqueous PTH conversion layers of seryl residues 15, 17, 18, 19, and 35 were subjected to phosphorus analysis to determine the presence or absence of a phosphorylated residue at that site (3).

Results

Fig. 1 shows disc gel electrophoretic patterns at pH 9.6 and 4 M urea for β-C, its derivatives, TFA β-C, and trypsin cleaved TFA β-C_{Fg}. Masking the lysine residues in β-C with TFA results in a homogeneous product with a mobility faster than that of β-C. Cleavage of TFA β-C with trypsin yields the large fragment of β-C containing two major and one minor band with mobilities similar to that of TFA β-C. On removal of TFA, the resulting β-C_{Fg}

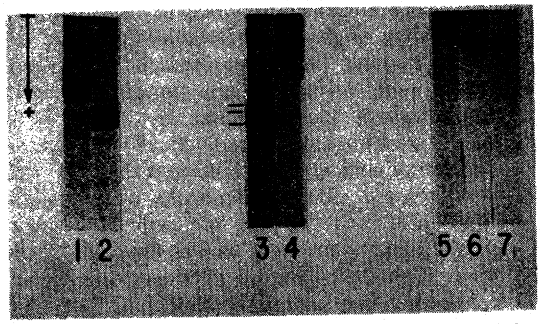

FIG. 1. Disc gel electrophoretic patterns pH 9.6, 4 M urea: (1) β-C; (2) TFA β-C; (3) Trypsin cleaved TFA β-C_{Fg} (the three lines are drawn to indicate the three bands which appear as one zone in the photograph), (4) β-C_{Fg}; (5) β-C_{Fg} I; (6) β-C_{Fg} II; and (7) β-C_{Fg} III.

[2] Reference to brand or firm name does not constitute endorsement by the U.S. Department of Agriculture over others of a similar nature not mentioned.

[3] Fg denotes fragment.

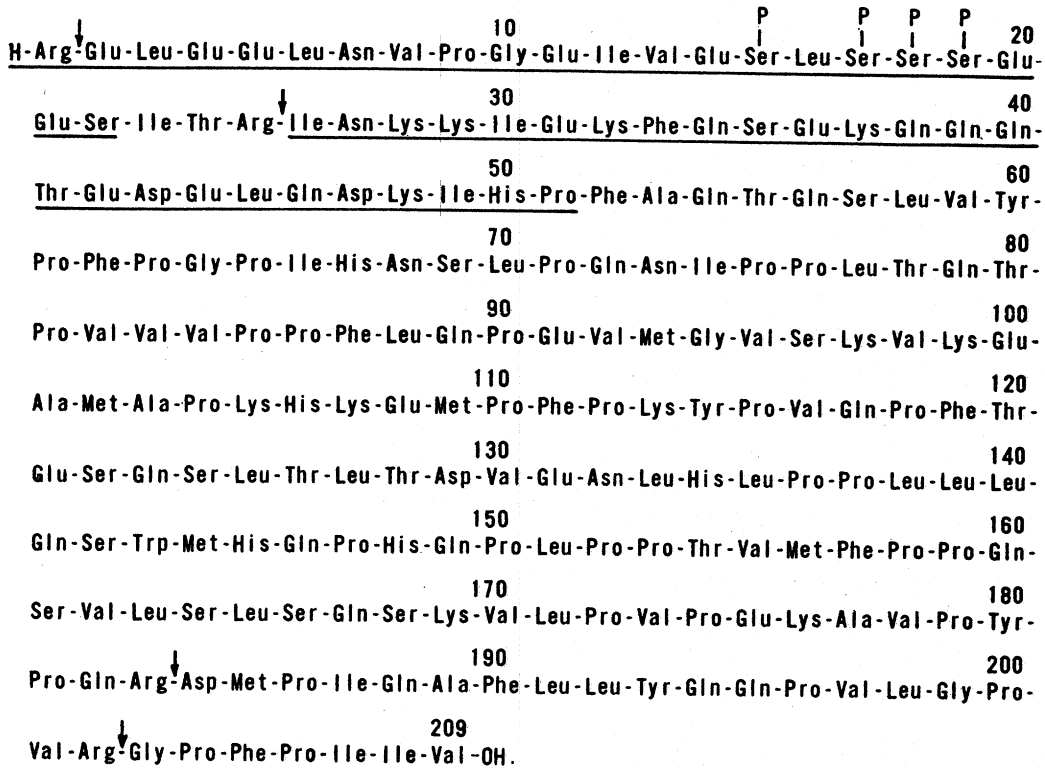

FIG. 2. Primary structure of β-casein C according to Grosclaude et al. (4) (reproduced by permission). The amino acid sequences determined in the present study are underlined.

shows three zones of relatively slow mobility (Fig. 1) designated β-C_{Fg} I, II, and III. Patterns of the isolated peptides β-C_{Fg} I, II, and III are also in Fig. 1.

Fig. 2 shows the primary structure of β-casein C as reported by Grosclaude et al. (4). Trypsin should cleave TFA β-C at the four arginines as indicated by arrows at residues 1, 25, 183, and 202, yielding one large peptide containing 158 residues (26 to 184), three other small peptides with 25, 19, and 7 residues and one arginine.

A comparison of the amino acid composition for β-C, residues 26 to 184, shows excellent agreement with the large peptide β-C_{Fg} I isolated after trypsin cleavage of TFA β-C (Table 1). The one residue differences for glutamic acid and proline are within experimental error and the lower value for serine is probably due to hydrolytic destruction. The lower values for valine and isoleucine probably result from incomplete hydrolysis of the isoleucine and valine bonds present in β-C. β-C_{Fg} II differs from I by an extra serine and glycine, neither of which can account for the difference in their electrophoretic mobility. The smallest fraction, β-C_{Fg} III, shows a slightly lower value for arginine and one less lysine. The reason for the differences in composition of the three fragments is not apparent.

The N-terminal sequence of β-C, residues 1 to 23, underlined in Fig. 2, was identical to that reported by Grosclaude et al. (4). This includes residues 15, 17, 18, and 19, which are phosphoserines. Phosphorus was determined on the aqueous layers from the PTH conversion of these residues.

The N-terminal sequence of β-C_{Fg} I agrees with the 26 residues, 26 to 52, shown in Fig. 2 for β-C. Phosphorus determination was negative for serine residue 35. Thus, residue 35 is not phosphorylated.

Discussion

Ethyl thioltrifluoroacetate is effective in modifying β-casein so that a large fragment (residues 26 to 184, Fig. 1) can be isolated from β-C for sequence studies.

The investigators at Jouy have determined the complete amino acid sequence of β-casein

TABLE 1. Comparison of the amino acid residues in a large fragment of β-C with the isolated fragments β-C_{Fg} I, II, and III.[a]

Amino acid	β-C Residues 26 to 184	β-C_{Fg} I[b]	β-C_{Fg} II[b]	β-C_{Fg} III[c]
Lys	12	12.0	12.1	11.2
His	6	5.7	5.8	5.6
Arg	1	0.9	0.9	0.6
Asp	7	7.0	7.2	7.2
Thr	8	7.5	7.6	7.5
Ser	11	9.8	11.0	9.6
Glu	28	29.0	28.5	27.3
Pro	28	29.0	27.1	25.0
Gly	2	2.1	3.2	2.2
Ala	4	4.1	4.4	3.8
Cys	...	...	...	...
Val	14	13.7	13.4	12.6
Met	5	5.0	4.8	4.6
Ile	5	4.4	4.3	4.3
Leu	16	16.6	16.2	15.5
Tyr	3	2.9	3.0	2.6
Phe	7	7.0	6.8	6.8
Trp	1	...[d]	...[d]	...[d]

[a] Residues calculated from molar ratios based on Asp = 7, Phe = 7.
[b] Average of two determinations.
[c] One determination.
[d] Tryptophan was not determined.

A^2 (11) and the sequence of the peptide in which substitution, glu/lys, which is characteristic of β-C, occurs (4). Thus, peptide 33 to 38 of β-C, Fig. 1, was sequenced and identified as the peptide in which residue 37 is lysine and in which serine residue 35 is free of phosphate. They found too that the peptide, residues 59 to 69 in β-C has at residue 67 histidine rather than the proline found in β-A^2. Results of our analysis on β-C and β-C_{Fg}, residues 1 to 23 and 26 to 52 (underlined in Fig. 1) agree with the sequence reported by Grosclaude et al. (4).

We also find that the serine at residue 35 is not phosphorylated and that residue 37 is lysine and not glutamic acid which is present in the other β-casein polymorphs. If γ-casein results from a specific enzymatic cleavage of β-casein at residue 29, it appears that the lysine substitution for glutamic acid in β-C at residue 37 prevents the enzymatic phosphorylation of serine at 35, and these changes, in turn, modify the environment around residue 29 so that cleavage to produce γ-C does not occur. Mercier et al. (9) proposed a mechanism for the enzymatic phosphorylation of caseins which requires that a second residue following a serine or threonine must be a phosphoserine or glutamic acid residue before phosphorylation can occur.

Alternatively, if the change in polarity around residue 35 and 37 due to an amino acid substitution in β-C compared to β-A^2 does not affect the lability of the bond between 28 to 29, it is possible that the lysine at 37 substituted for glutamic acid in β-C is sensitive to a trypsin-like enzyme and cleavage occurs not only after lysine 28 where γ-casein should be produced, but also after lysine 37, so that no γ-casein would be found. On disc gel electrophoresis (pH 9.6) of casein samples containing β-casein C, a zone of slightly slower mobility than that of TS-casein C is always seen. This band is absent in caseins which do not contain β-C. Whether this zone represents a new fragment of β-C is unknown.

References

(1) Goldberger, R. F. 1967. Trifluoroacetylation of ε-amino groups. Method. Enzymol. 11: 317.

(2) Gordon, W. G., M. L. Groves, R. Greenberg, S. B. Jones, E. B. Kalan, R. F. Peterson, and R. E. Townend. 1972. Probable identification of γ-, TS-, R-, and S-caseins as fragments of β-casein. J. Dairy Sci. 55:261.

(3) Greenberg, R., M. L. Groves, and R. F. Peterson. 1974. Human casein: Amino terminal sequence and location of phosphate groups. Fed. Proc. Fed. Amer. Soc. Exp. Biol. 33: 1563.

(4) Grosclaude, F., M.-F. Mahé, J.-C. Mercier, and B. Ribadeau Dumas. 1972. Caractérisation des variants génétiques des caséine α_{s1} et β bovines. Eur. J. Biochem. 26:328.

(5) Groves, M. L., and W. G. Gordon. 1969. Evidence from amino acid analysis for a relationship in the biosynthesis of γ- and β-caseins. Biochim. Biophys. Acta 194:421.

(6) Groves, M. L., W. G. Gordon, E. B. Kalan, and S. B. Jones. 1972. Composition of bovine γ-caseins A^1 and A^3, and further evidence for a relationship in biosynthesis of γ- and β-caseins. J. Dairy Sci. 55:1041.

(7) Groves, M. L., W. G. Gordon, E. B. Kalan, and S. B. Jones. 1973. TS-A^2, TS-B, R-, and S-caseins: Their isolation, composition, and relationship to the β- and γ-casein polymorphs A^2 and B. J. Dairy Sci. 56:558.

(8) Groves, M. L., and C. A. Kiddy. 1968. Polymorphism of γ-casein in cow's milk. Arch. Biochem. Biophys. 126:188.

(9) Mercier, J.-C., F. Grosclaude and B. Ribadeau Dumas. 1972. Primary structure of bovine caseins. A review. Milchwissenschaft 27:402.

(10) Moore, S., and W. H. Stein. 1963. Chromatographic determination of amino acids by the use of automatic recording equipment. Methods. Enzymol. 6:819.

(11) Ribadeau Dumas, B., G. Brignon, F. Grosclaude, and J.-C. Mercier. 1972. Structure primaire de la caséine β-bovine. Sequence complète. Eur. J. Biochem. 25:505.

(12) Smithies, O., D. Gibson, E. M. Fanning, R. M. Goodfliesh, J. G. Gilman, and D. L. Ballantyne. 1971. Quantitative procedures for use with the Edman-Begg sequenator. Partial sequences of two unusual immunoglobulin light chains, Rzf and Sac. Biochemistry 10:4912.